SUB | NIEDERSÄCHSISCHE STAATS- UND UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

The ASCH Project and the Modeling of Linked Data Compliant Provenance Information

Stefanie Rühle (SUB Göttingen)
Opening up the collection – reuse and publishing
07.06.2016

Overview

- About the project and the collections
- How to start
- Modelling

About the Project

- Purpose and Scope:
  - developing a metadata model for contextualising heterogenous objects from collections
  - with special focus on provenance information

- Funding
  - by Deutsche Forschungsgemeinschaft
  - 01.09. 2014 – 31.08. 2017

- Project Lead:
  - Göttingen State and University Library
  - Institute of Social and Cultural Anthropology at the University of Göttingen

- Contact
  - Susanne Al-Eryani (mailto: al-eryani(a)sub.uni-goettingen.de)
  - Gudrun Bucher (mailto: gudrun.bucher(a)sub.uni-goettingen.de)
  - Jürgen Dönitz (mailto: doenitz(a)sub.uni-goettingen.de)

Description of Provenance

- „… descriptions of the entities and activities involved in producing and delivering or otherwise influencing a given object.“

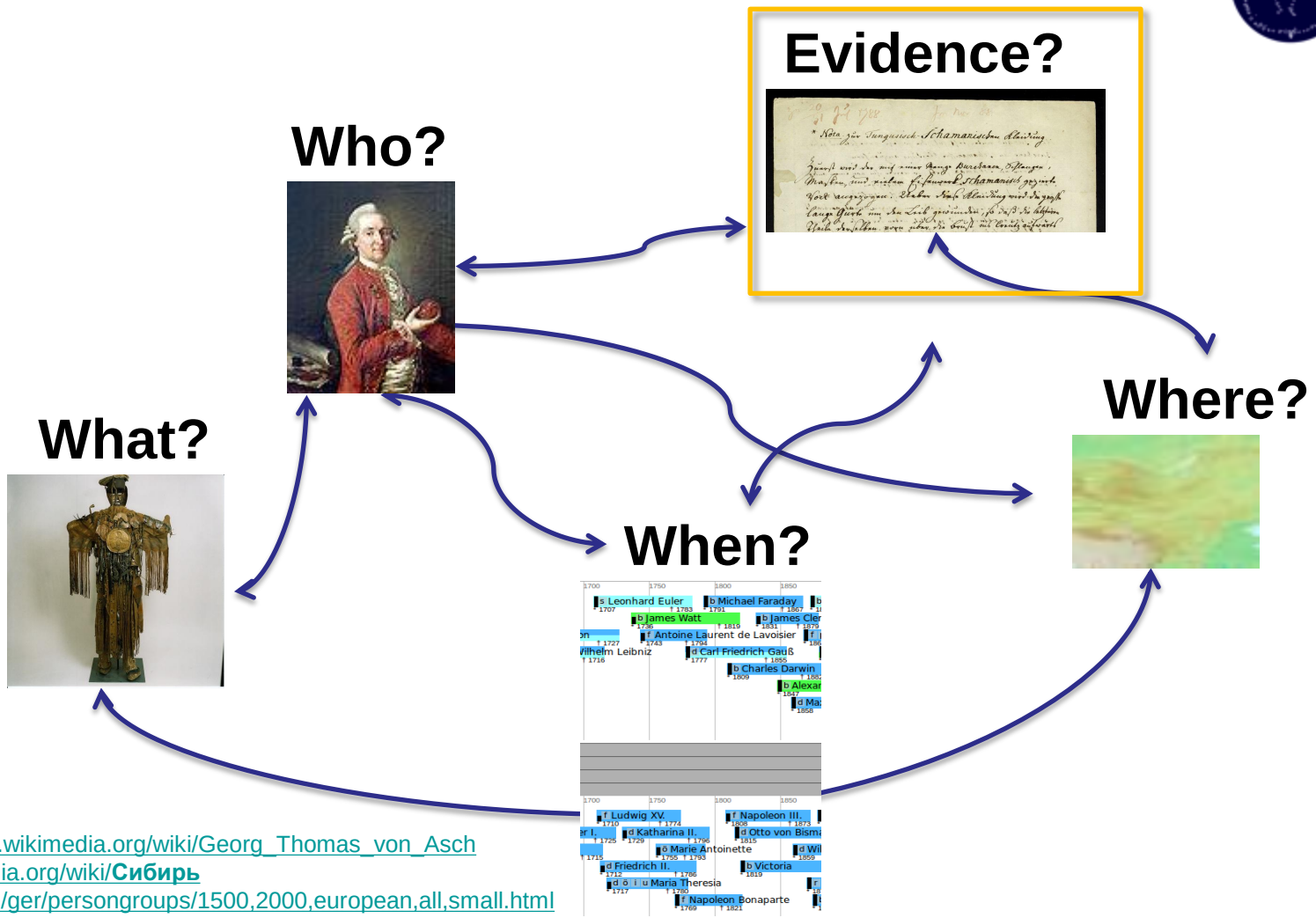
  Source: https://www.w3.org/TR/prov-overview/

- Description of the events in the lifecycle of an object
  - LIDO (Lightweight Information Describing Objects)
  - CIDOC-CRM (Conceptual Reference Model of the International Committee for Documentation of the International Council of Museums)
  - W3C PROV (a specification to express provenance records)

- Involved entities: what, who, when, where

- Evidence verifying statements

Contextualization of Provenance Information



Source: http://commons.wikimedia.org/wiki/Georg_Thomas_von_Asch
http://commons.wikimedia.org/wiki/**Сибирь**
http://www.vistorica.com/ger/persongroups/1500,2000,european,all,small.html

Cultural and Scientific Heritage in the WWW

- The vision
  - A one-stop shop for access to millions of digital cultural and scientific objects
  - Interlinking the results of different projects, institutions and disciplines

- The need
  - well-structured consistent description of objects using metadata standards

- The reality
  - Metadata descriptions are aligned to the requirements of specific projects, in-house rules, legacy data, community agreements, …
  - Metadata standards used by cultural and scientific institutions sometimes differ widely in their model and structures

And most time it makes sense!

„Framework“

- Uses the least common denominator to bring the data together
- Allows different types of metadata descriptions
- Steps in this direction back to the 90s
  - Dublin Core Dumb-Down Principle: „… a principled way of viewing a complex metadata description through the lens of a simpler representation …“

    (source: http://wiki.dublincore.org/index.php/Glossary/Dumb-Down_Principle)

  - RDF (Resource Description Framework): a most generic framework modelling the relations between resources
    - Mandatory: every described resource must be identified by a unique http-Identifier
    - Freedom: use whatever metadata standard you like as long as the terms are identified by http-URIs

Why Provenance „von Asch“?

Georg Thomas von Asch (1729-1807)

portraiture by Kyrill Golowatschewski,
Oil on canvas, 86 x 70 cm,
donation of 1780

Heterogeneity of the Collection

- Manuscripts, books, maps (Göttingen State and University Library)
- Coins and medals (Institute of Archaeology)
- Anthropological items from different regions of the world (Institute of Social and Cultural Anthropology)
- Rocks, minerals, meteorite, fossils (Geoscience Centre)
- Engravings and portraits (Arts Collection)
- Human sculls (Centre for Anatomy)
- Mounted animals (Museum of Zoology)
- Seeds (?)

Heterogeneity of Provenance

Kjachta, Siberia

Russian America

Turkey

Black See Region

St. Petersburg

???

Institute of Arts Archaeologie

Anatomy

Academisches Museum Göttingen (1773)

Zoology

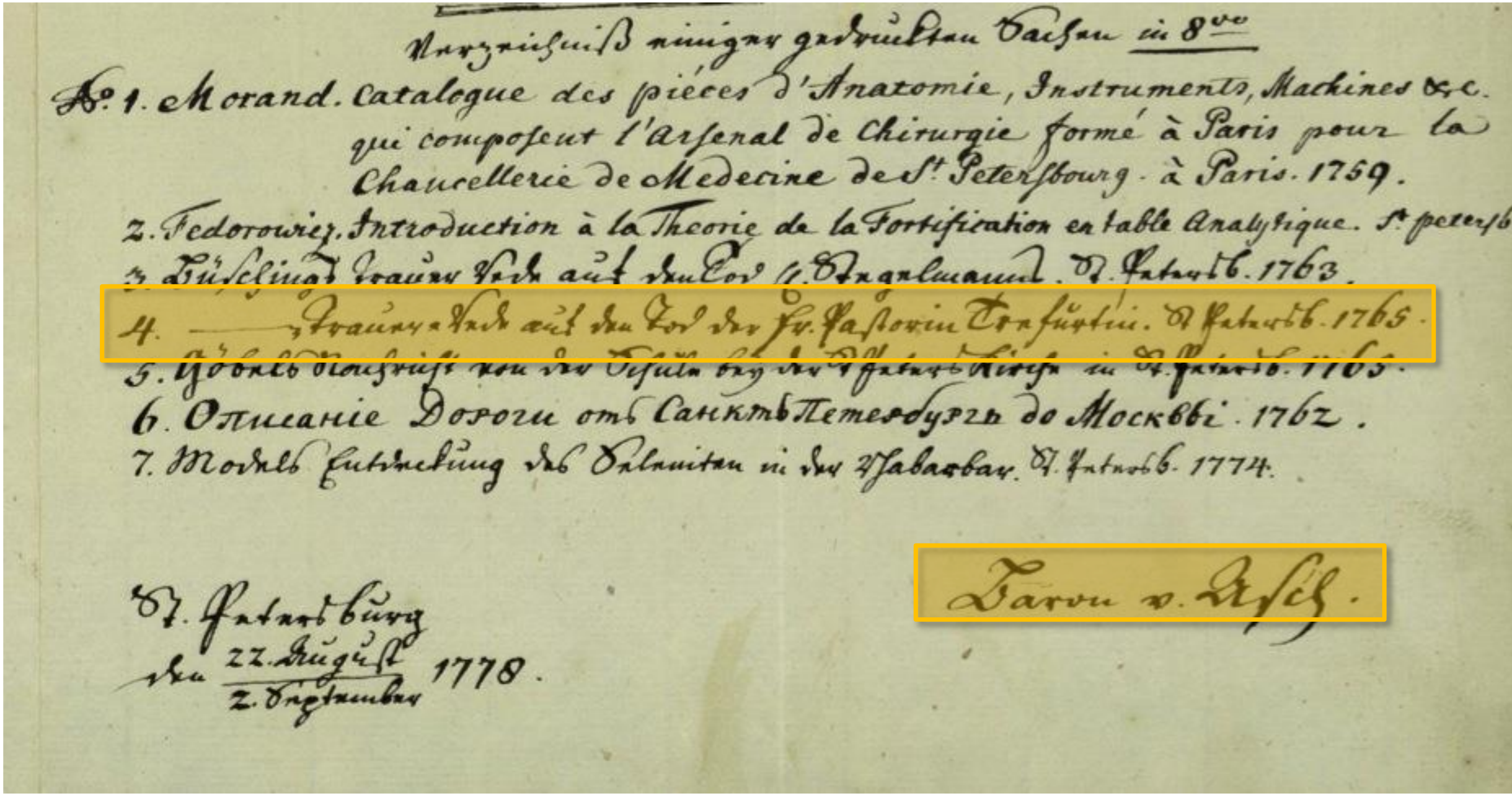
Geoscience

University Library

Cultural Anthropolgy

Von Asch Evidence

- Letters and listings from von Asch
- Letters by others about the items von Asch sent to Göttingen
- Lists and inventories of the university
- Notes about single items
- Labels (written by von Asch himself or later by Blumenbach et al. in Göttingen)


Verzeichniß einiger gedruckten Sachen in 8vo

No. 1. Morand. Catalogue des pièces d'Anatomie, Instruments, Machines &c. qui composent l'Arsenal de Chirurgie formé à Paris pour la Chancellerie de Medecine de St. Petersbourg. à Paris. 1759.

2. Fedorowiez. Introduction à la Theorie de la Fortification en table Analytique. St. petersb.

3. Büschings TrauerRede auf den Tod d. Engelmann. St. Petersb. 1763.

4. —— TrauerRede auf den Tod der Fr. Pastorin Trefurtin. St. Petersb. 1765.

5. Göbels Denkschrift von der Schule bey der Petri Kirche in St. Petersb. 1765.

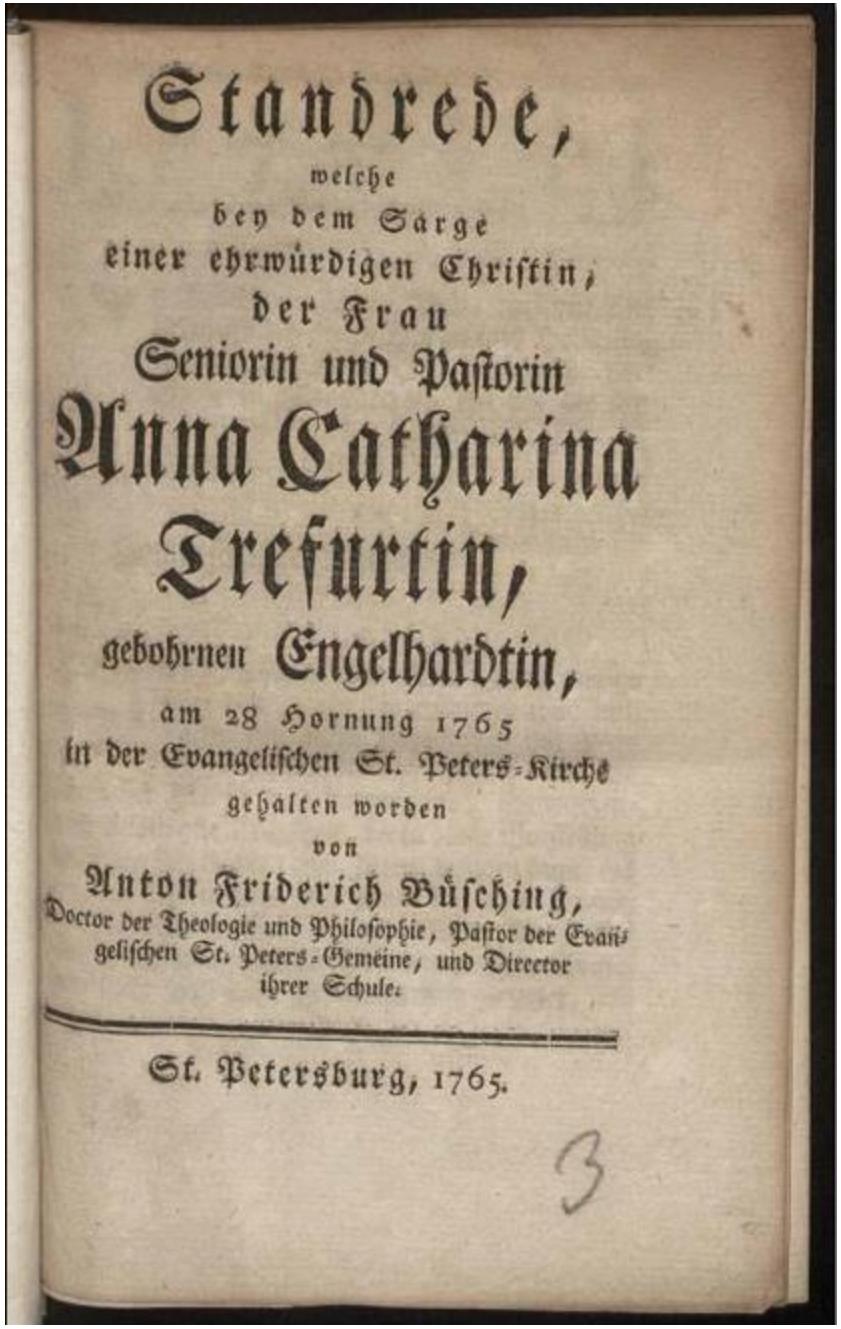
6. Описание Дороги отъ Санктъпетербурга до Москвы. 1762.

7. Models Entdeckung des Seleniten in der Rhabarbar. St. Petersb. 1774.

St. Petersburg den 22. August / 2. September 1778.

Baron v. Asch.

____ [Büschings] TrauerRede auf den Tod der Fr. Pastorin Trefurtin. St. Ptetersb. 1765



Standrede,
welche
bey dem Sarge
einer ehrwürdigen Christin,
der Frau
Seniorin und Pastorin
Anna Catharina
Trefurtin,
gebohrnen Engelhardtin,
am 28 Hornung 1765
in der Evangelischen St. Peters-Kirche
gehalten worden
von
Anton Friderich Büsching,
Doctor der Theologie und Philosophie, Pastor der Evangelischen St. Peters-Gemeine, und Director ihrer Schule.

St. Petersburg, 1765.

3

Title page of the book, now part of the stock of the Göttingen State and University Library.

Digital Copy:

https://www.deutsche-digitale-bibliothek.de/item/6IOEUJNCLFUMCADYCPFRXIL7V4AS66X6

zu 20/21 Jul 1788 Jm Nov 88

* Nota zur Tungusisch-Schamanischen Kleidung

Zuerst wird der mit einer Menge Burchanen, Schlangen, Masken, und vielem Eisenwerk schamanisch gezierte Rock angezogen. Ueber diese Kleidung wird die grosse lange Gurte um den Leib gewunden, so daß die letzteren Theile derselben vorn über die Brust ins Creutz aufwärts

...scription wird auch zur rechten Seite am Rocke befestigt. Der Schaman giebt vor, daß sie sein Diplom sey.

Die rasende Stellung, könnte ungefehr so nachgeahmet werden, wie sie der Schaman Tab. 68. (vid. Vorstellung tibetischer Stationen) dargestellt ist.

Nur die Zauber Trummel u. Schlägel fehlt, noch dazu.

Unter der Menge von Burchanen, Schlangen, Glocken, und so vielen eisernen Anhängseln am Schamanischen Kleide, bemerkt man auch ledernen Säckelchen, worinn Stücke Fleisch aufbewahrt ist.

Bärs

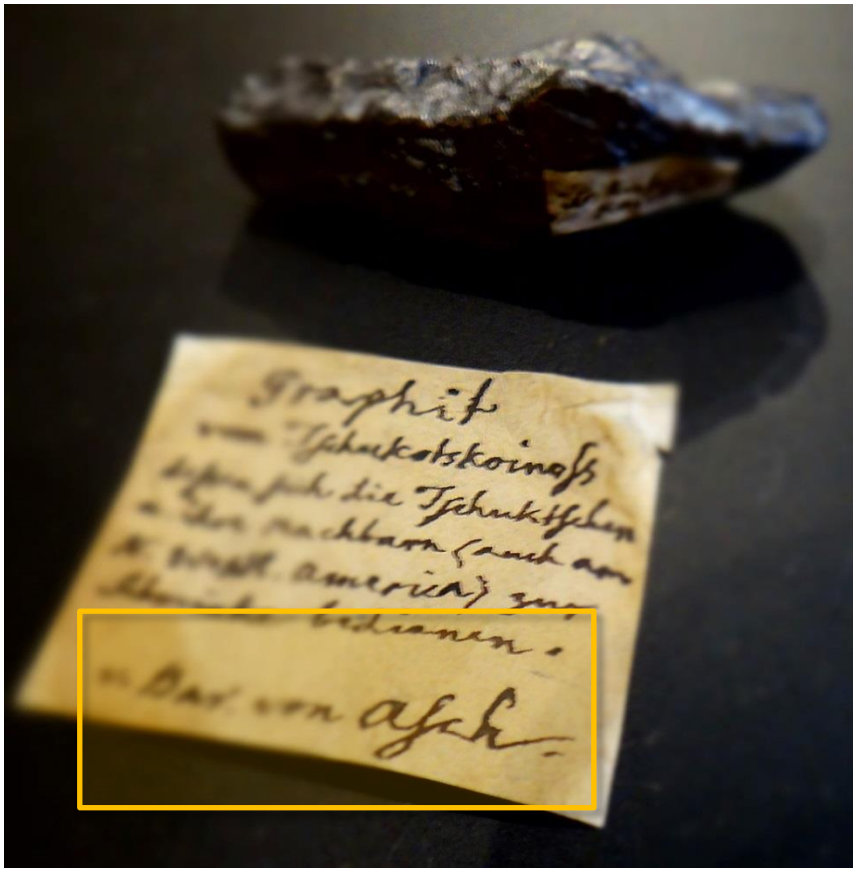
Ethnographic Collection

Shamans costume back, Siberia,

Tungus (Evenks),

As 957

Geoscience Centre: Historical Collections



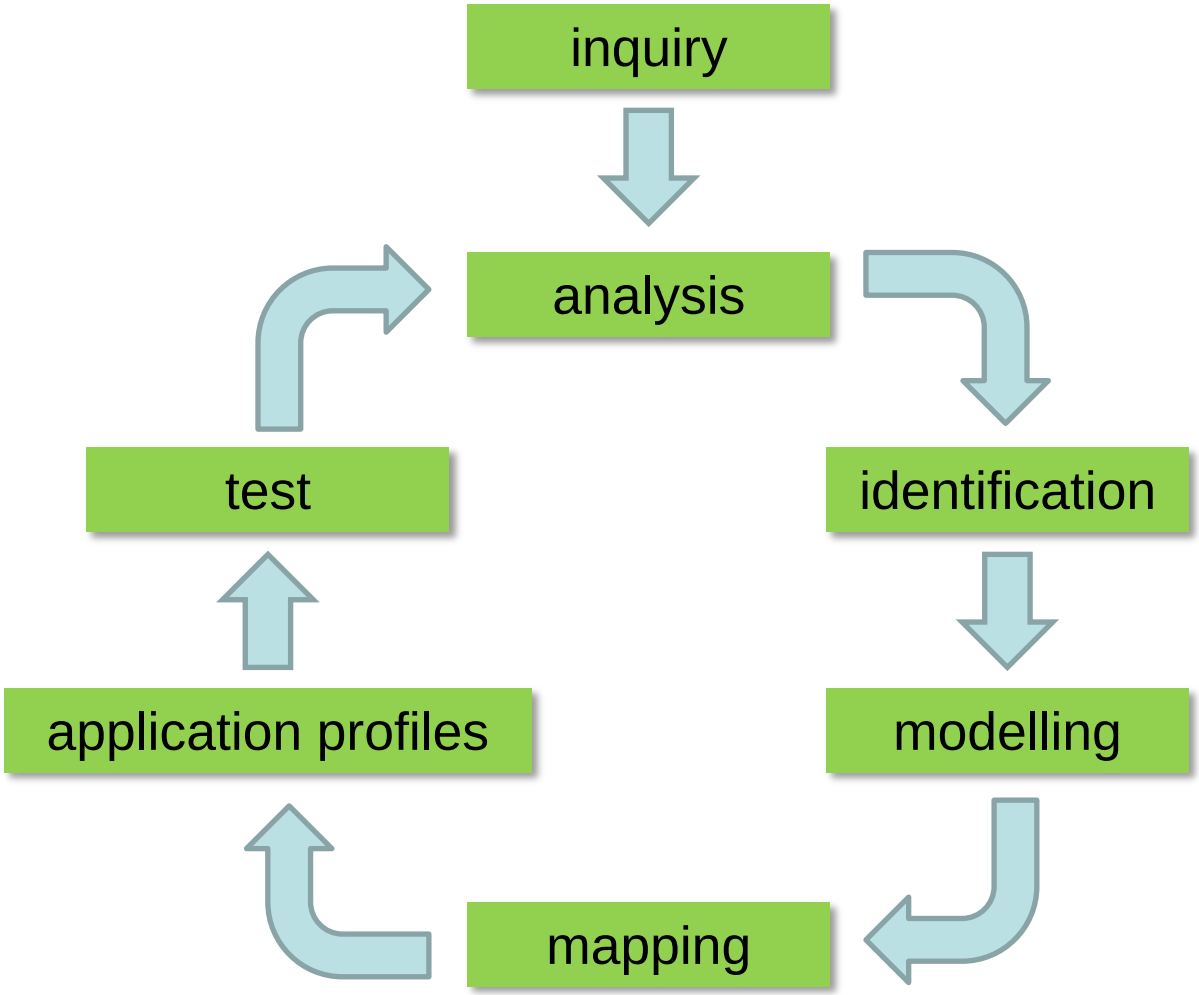
Graphite; Geoscience Museum;

with original labeling – von Asch gives some information on the graphit‘s use

Overview

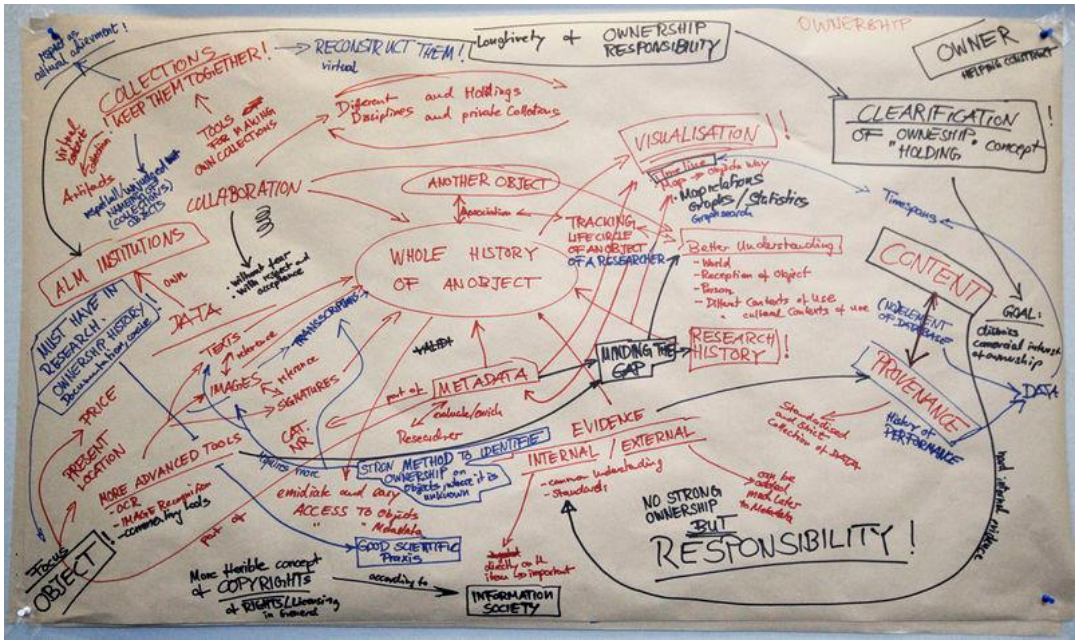
- About the project and the collections
- How to start
- Modelling

7 Steps to your Metadata Model



1: Inquiry

- **Workshop with experts from humanities and natural sciences**
- across domains
  - Museums
  - Libraries
  - Archives
- across disciplines
  - Anatomy
  - Archaeology
  - Library Science
  - Botany
  - Cultural Anthropology
  - Geology / Mineralogy
  - Computer Science
  - History of Art
  - Musicology
  - Zoology
- **Interviews**



Change of ownership / Chain of custody – Visions
http://asch.wiki.gwdg.de/index.php/Workshop_2015

2. Analysis

ID	Case Study	Scenario ID	Scenario
	readings about the expeditions.	Scenario 37	A user needs evidence related to an event.
CS 5	Fr…to identify origin and age o…ns the University got from Ba… Asch. He us… correspondence and record… von Asch as a primary resourc… his studies. He …ecially interested in a tetradrachm showing Apollo that according to the letters fro… Asch was coined in… inf…on about the origin… sim…ces in other colle… for other collections with co… coins collected at the Crime… place. Therefore he needs t… expeditions, recordings and … with collectors, etc.). He thi… better insight regarding the…	Scenario 36	A user needs the evidence related to an item.
CS 6	Adora learns from the regist… von Asch sent a collection of… to be lost since the beginnin… to the collection. She enters … materials like letters and list… the inventories listing the it… find similar items in other c… Asch collection. Therefore s… descriptions of prints in othe…		
CS 7	Jürgen is working on a study … especially for publications fr… influenced the war against p… gets a long list of publicatio… know when and where thes… owned or who had access to … with reports on immunizatio…		
CS 8	Kerstin is working on a comp… Collection she finds an inter…		

Use Case	ID	Scenario	Actor	Goal	Requirement ID	Requirement
Scenarios related to information about the event/activities the items of a collection were involved						
UC 3 Information about the history/lifecycle of resources	Scenario 22	A user needs information about all events in the lifecycle of an item during a certain time span.	user	Find events related to an item during a certain time span	[Requirement 20]	Item descriptions must be interlinked with 1-n events (creation, modification, collection, etc.) in the lifecycle of the item.
					[Requirement 27]	An event in the lifecycle of an item must be related to 0-n date information.
UC 3 Information about the history/lifecycle of resources	Scenario 23	A user needs information about items that were present during an event (e.g. collected items, used items, created items, etc.).	user	Find items related to a certain event.	[Requirement 23]	An event in the lifecycle of an item must be related to 1-n items.
UC 3 Information about the history/lifecycle of resources	Scenario 24	A user is searching for items of the same price related to acquisition events.	user	Find events of the same type having statements with similar values	[Requirement 20]	Item descriptions must be interlinked with 1-n events (creation, modification, collection, etc.) in the lifecycle of the item.
					[Requirement 8]	The nature of a resource must be described using 1-n controlled values.
					[Requirement 1]	Resource descriptions must be machine readable.
UC 3 Information aobut history/lifecycle of resources	Scenario 25	A user wants to know where an event happened.	user	Find places related to an event	[Requirement 26]	An event in the lifecycle of an item must be related to 0-n places.
					[Requirement 39]	Places must be identified by identifiers that are unique, machine readable and persistent.
					[Requirement 41]	Place descriptions can include 1-n place appellations
UC 3 Information about history/lifecycle of resources	Scenario 26	A user wants to know more about places related to an item.	user	Get information about places related to a certain item.	[Requirement 9]	Resources can be interlinked with 0 - n other resources.
					[Requirement 3]	A resource description must provide all attributes necessary to identify the resource and distinguish it from other resources of the same type.
UC 3 Information about history/lifecycle of resources	Scenario 27	A user is searching for items that have been collected in a certain area.	user	Find items related to the same type of events happening at a certain place.	[Requirement 20]	Item descriptions must be interlinked with 1-n events (creation, modification, collection, etc.) in the lifecycle of the item.
					[Requirement 22]	Event descriptions must show the nature of the relation between items, agents, places and/or time.
					[Requirement 26]	An event in the lifecycle of an item must be related to 0-n places.
					[Requirement 40]	Places must be identified by machine readable geo-information / geographic coordinates.

http://asch.wiki.gwdg.de/images/2/28/20150928_ASCH_UseCases_final.xls

Use Cases

- UC 1 Information about resources
- UC 2 Identification of resources
- UC 3 Information about the history/lifecycle of resources
- UC 4 Information about change of use and reception of resources
- UC 5 Proof of information by evidence
- UC 6 Reliability of statements
- UC 7 Access to resources
- UC 8 Reuse of data

http://asch.wiki.gwdg.de/index.php/Use_Cases

Requirements

ID	Requirement	Sub-Requirement of
	evidence = a resource proving the reliability of a statement about a resource	
[Requirement 15]	Evidences can be machine readable.	
[Requirement 16]	The nature of an evidence must be described by 1-n controlled values.	[Requirement 8]
[Requirement 17]	Information about resources in an evidence must be identified by unique identifiers.	
[Requirement 18]	Information about the nature of a resource in an evidence must be identified by 1-n controlled values.	
[Requirement 78]	Evidence descriptions must provide information about the accessibility and reusability of the resource.	
	event = an activity in the lifecycle of a resource	
[Requirement 79]	Event descriptions can be interlinked with evidence about the event (e.g. documents).	[Requirement 9]
[Requirement 21]	Event descriptions must be machine readable.	[Requirement 11]
[Requirement 22]	Event descriptions must show the nature of the relation between items, agents, places and/or time.	
[Requirement 23]	An event in the lifecycle of an item must be related to 1-n items.	[Requirement 9]
[Requirement 24]	An event in the lifecycle of an item must be related to 0-n agents.	[Requirement 9]
[Requirement 25]	The function/role of an agent during an event must be specified using controlled values (e.g. creator, dealer, collector).	[Requirement 11
[Requirement 26]	An event in the lifecycle of an item must be related to 0-n places.	[Requirement 9]
[Requirement 27]	An event in the lifecycle of an item must be related to 0-n date information.	[Requirement 9]
	time = a time span related to a resource via an acitivity or as a topic	
[Requirement 29]	Date and time indication must be machine readable.	[Requirement 1]
[Requirement 30]	Date information must be compliant to date type standards.	[Requirement 64]

Overview

- About the project and the collections
- How to start
- Modelling

3: Identification of relevant classes and relations between these classes

- Classes
  - resource
  - metadata set
  - item
  - **evidence**
  - **event**
  - time
  - agent
  - place
  - digital representation
  - collection
  - statement
  - holding
  - concept

Event

:Event
Type of Term: Class
Definition: An activity in the lifecycle of a resource.

:hasEvidence
Type of term: rdfs:Property
Definition: relates an event with resources proving the existence of the event
Domain: Event
Range: Evidence
Occurrence: 0-n
Subproperty of:

Evidence

:Evidence
Type of Term: Class
Definition: A resource proving the reliability of a statement about a resource
Subclass of:

4: Modeling the Use Cases

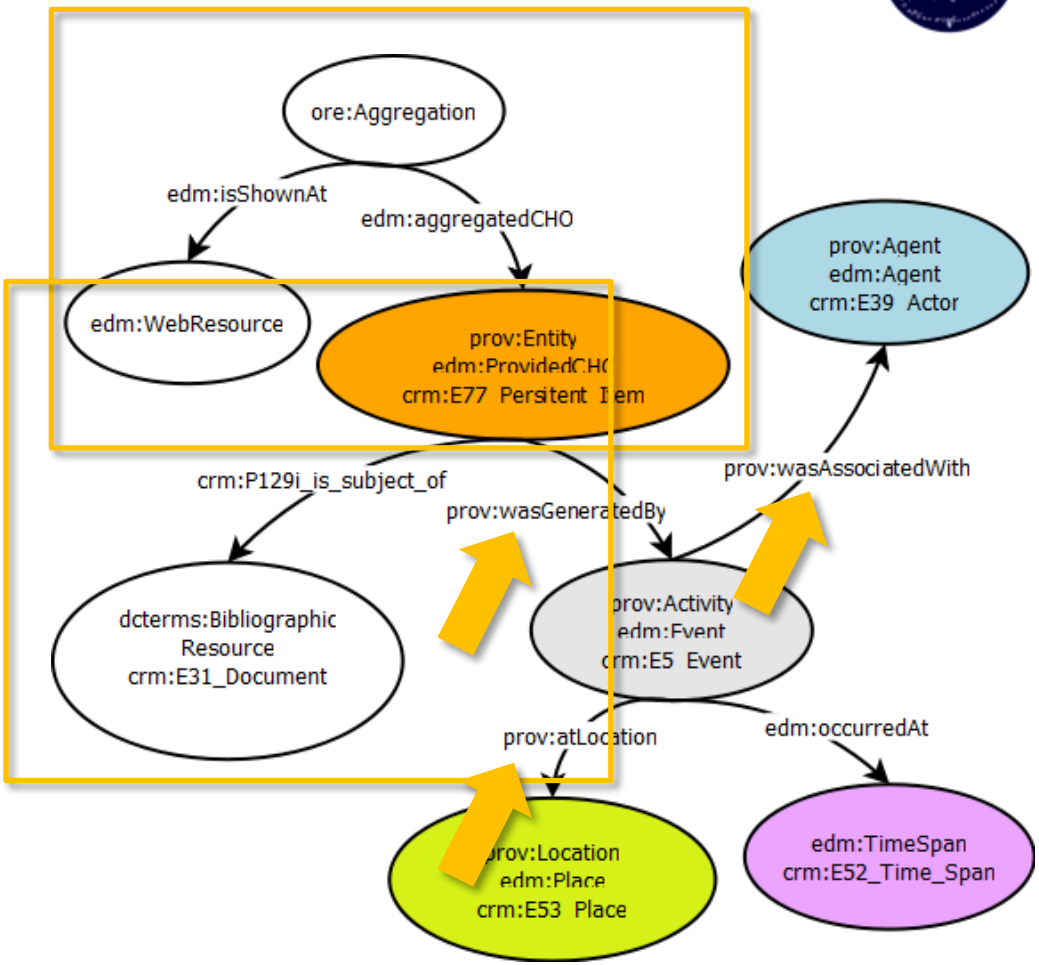
- UC 1 Information about resources
- UC 2 Identification of resources
- UC 3 Information about the history/lifecycle of resources
- UC 4 Information about change of use and reception of resources
- UC 5 Proof of information by evidence
- UC 6 Reliability of statements
- UC 7 Access to resources
- UC 8 Reuse of data

(**http://asch.wiki.gwdg.de/index.php/Use_Cases**)

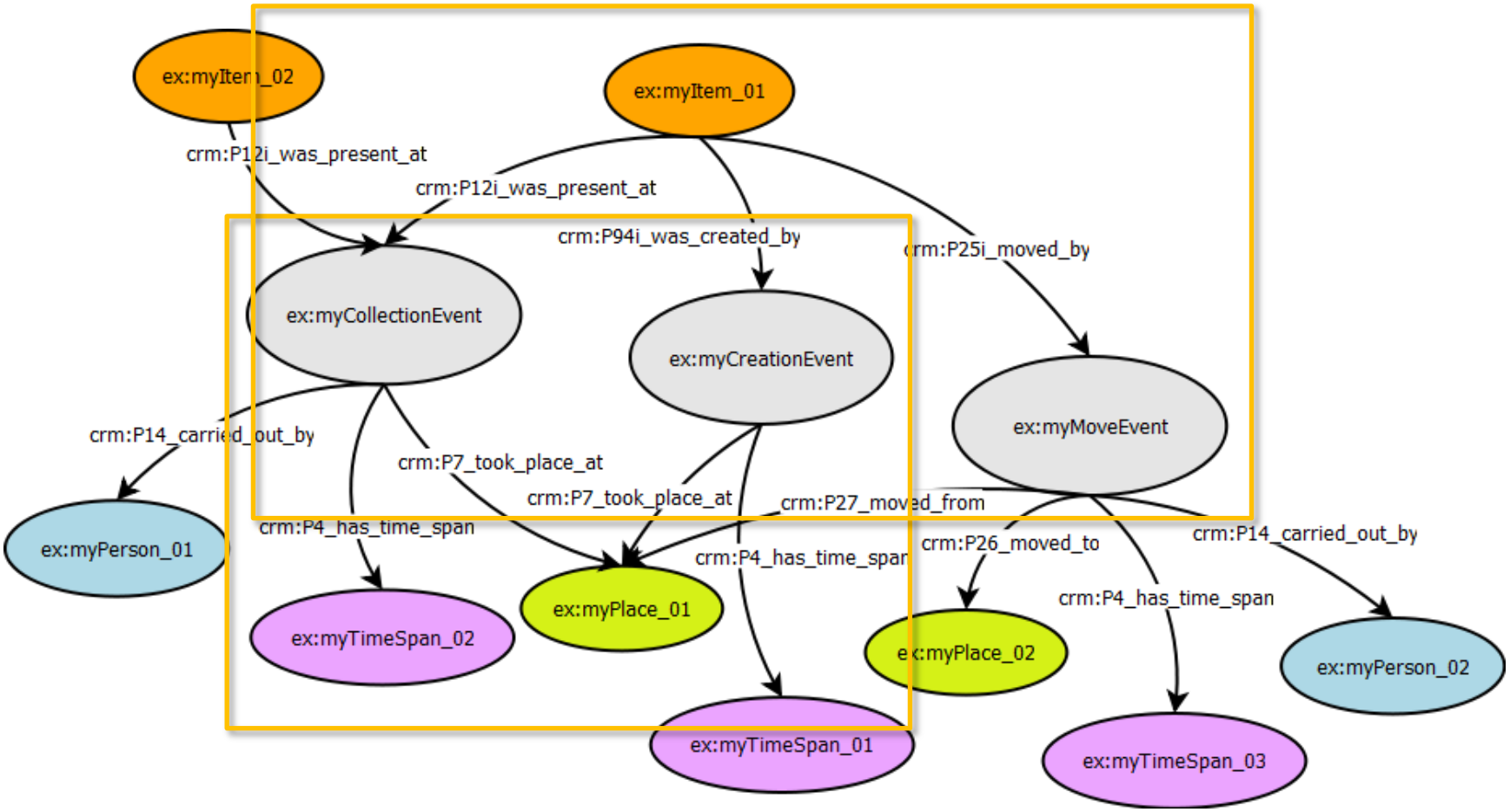
History/Lifecycle of a resource

- Developing a metadata framework to bring together provenance descriptions of different type

- Focusing on activities/events
  - W3C PROV as the backbone
  - Align it with other standards (e.g. EDM, CIDOC CRM)
  - Align it with digital representations
  - Align it with evidence

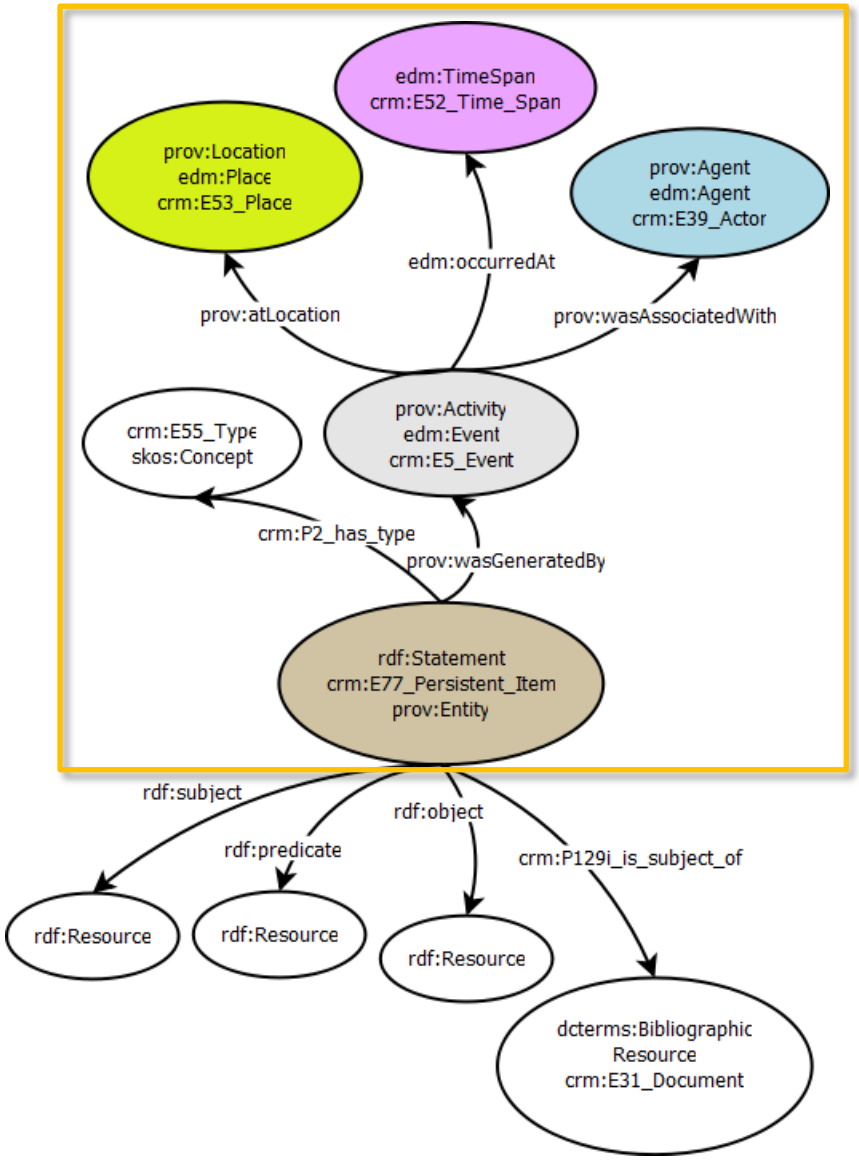


Event-based Description of the Lifecycle

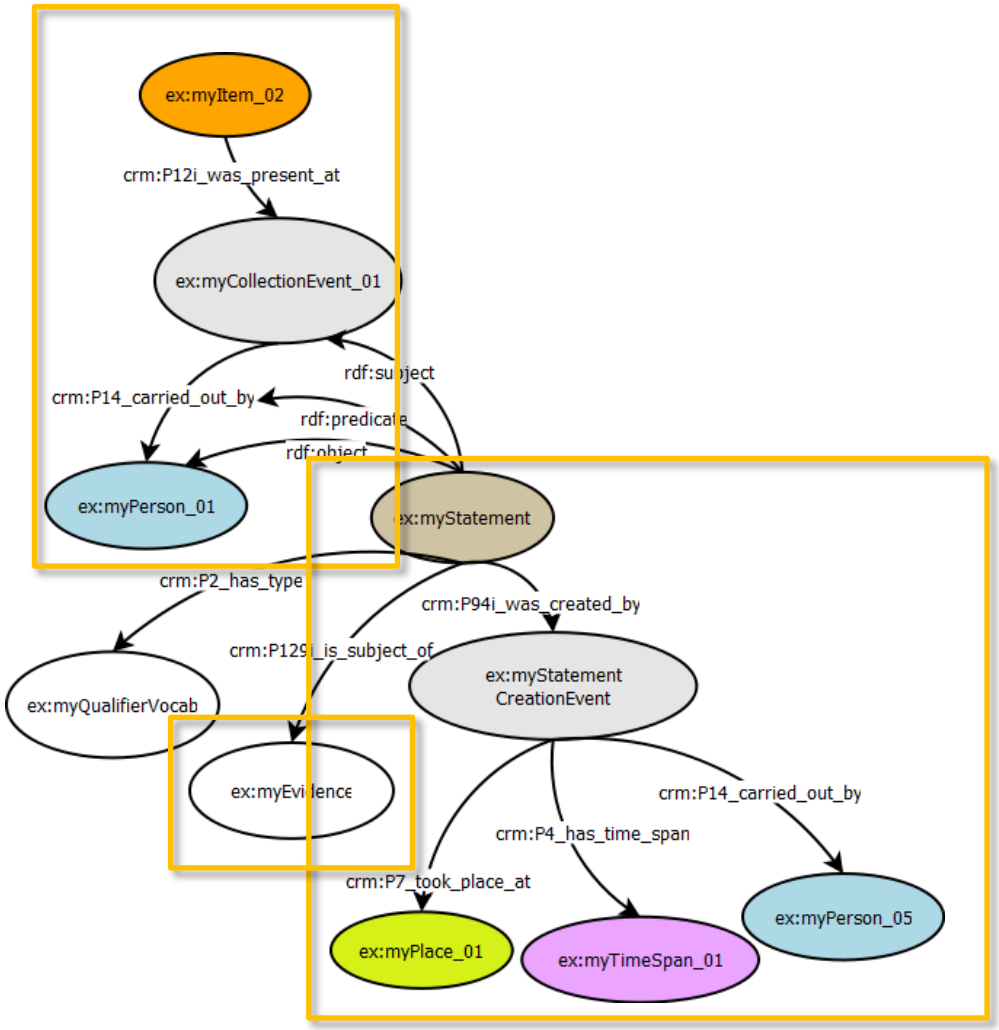


Statement Provenance

- Developing a metadata framework that allows to describe
  - Who made what statement when and where
  - How reliable a statement is
- Reuse of existing standards
  - Focusing on RDF -> Reification
  - Align it with other standards
  - Align it with evidence descriptions
- „Reification in this context means the expression of something in a language using the language, …“ (Source: https://www.w3.org/DesignIssues/Reify.html)



Description of Provenance Statements



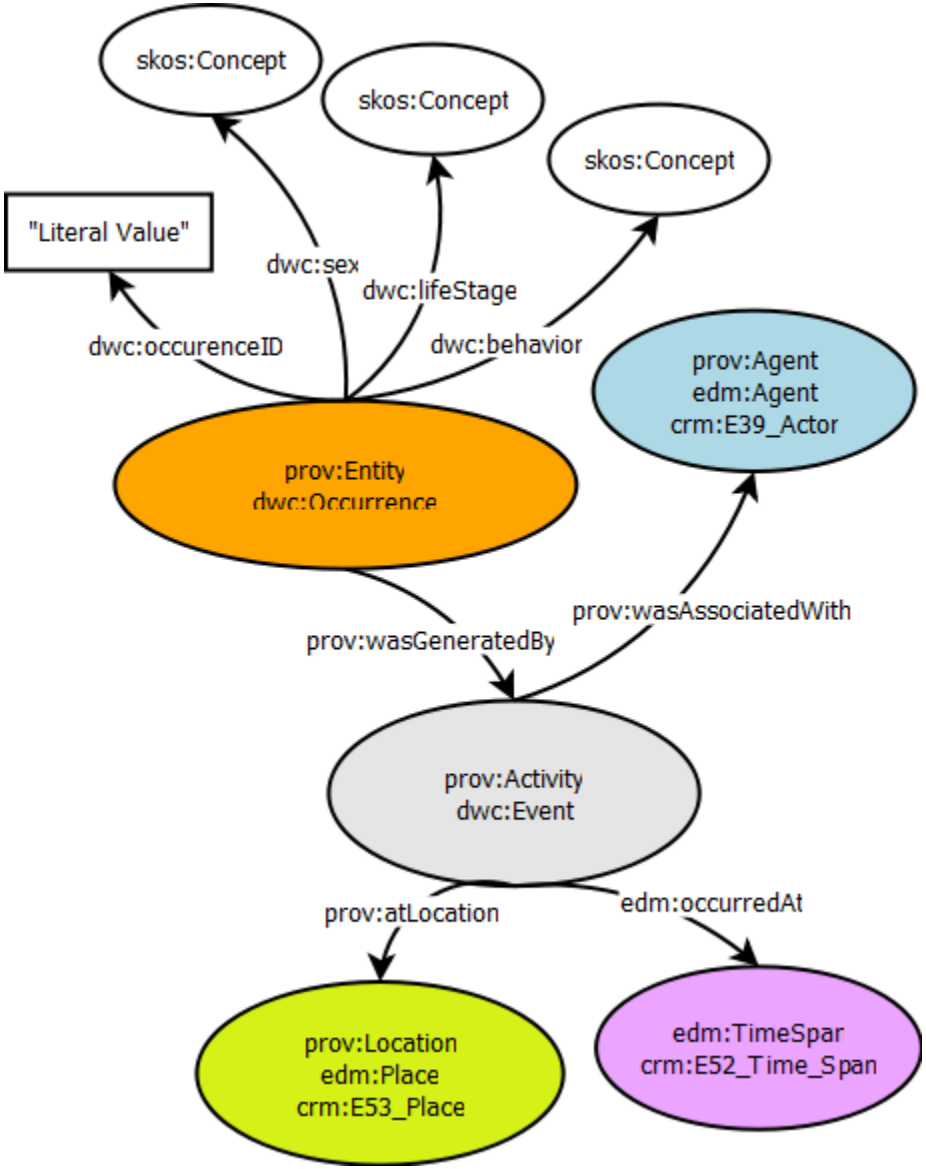
5: Check standards and requirements

[Requirement 78]	Evidence descriptions must provide information about the accessibility and reusability of the resource.	Property	P104 is subject to	This property links a particular E72 Legal Object to the instances of E30 Right to which it is subject.	crm:E73_Information_Object crm:P104_is_subject_to crm:E30_Rights	✓
Event	**An activity in the lifecycle of a resource**	**Entity**	**E5 Event**	**This class comprises changes of states in cultural, social or physical systems, regardless of scale, brought about by a series or group of coherent physical, cultural, technological or legal phenomena. Such changes of state will affect instances of E77 Persistent Item or its subclasses.**		✓
[Requirement 79]	Event descriptions can be interlinked with evidence about the event (e.g. documents)	Property	P129i is subject of	This property documents that an E89 Propositional Object has as subject an instance of E1 CRM Entity.	Crm:E5_Event crm:P129i_is_subject_of crm:E73_Information_Object	✓
[Requirement 21]	Event descriptions must be machine readable.			CRM supports a machine readable structured description of resources in TEXT, HTML, XML and RDF		✓
[Requirement 22]	Event descriptions must show the nature of the relation between items, agents, places and/or time.	Property	P2 has type	This property allows sub typing of CRM entities – a form of specialization – through the use of a terminological hierarchy, or	crm:E5_Information_Object crm:P2_has_type crm:E55_Type	✓

http://asch.wiki.gwdg.de/index.php/Metadata_Standards

6: Domain-specific Application Profiles

- Align provenance description with domain-specific standards
  - Darwin Core
  - MODS
  - nomisma.org
  - etc.
- Provide provenance information in different RDF-compliant standards
  - CIDOC-CRM
  - Darwin Core
  - LIDO
  - EDM
- Dumb-Down specific Properties and Classes



7: Test

- Test the model in different software environments:
  - WissKI → CIDOC-CRM AP
  - Semantic Media Wiki
  - OntoWiki
  - TextGrid
  - …

SUB | NIEDERSÄCHSISCHE STAATS- UND UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

Any questions?

Contact:
Susanne Al-Eryani (mailto: al-eryani(a)sub.uni-goettingen.de)
Gudrun Bucher (mailto: gudrun.bucher(a)sub.uni-goettingen.de)
Jürgen Dönitz (mailto: doenitz(a)sub.uni-goettingen.de)
Stefanie Rühle (mailto: sruehle(a)sub.uni-goettingen.de)

and more information at:
http://asch.wiki.gwdg.de/index.php/Main_Page